

Exercise Sheet 12

Submit until Tuesday, January 30 at **12:00pm (noon)**

Exercise 1 (14 points)

Write a program that reads a SPARQL query from a given file, translates it to a corresponding SQL query, and executes that query on a given SQLite3 database. The idea of the translation and the (simple) Python interface to SQLite3 were explained in the lecture; see also the code provided on the Wiki. Your program should support all SPARQL queries of the form

SELECT *<variables>* *WHERE* { *<triples>* }

where *<variables>* is a whitespace-separated list of variables, and *<triples>* is a dot-separated list of triples, where the three components of each triple (subject, predicate, object) and the dots are separated by whitespaces, and each component can be either a variable (e.g. *?x*) or a string literal (e.g. *"Jennifer Lopez"*). Also support the optional *LIMIT* *<k>* clause, which (if appended to the end of the query) restricts the output to (an arbitrary subset of) *k* results.

As usual, make sure that the test cases given in the TIP file on the Wiki are fulfilled. Also as usual, feel free to adapt the exact syntax of the test cases to your particular implementation. This is especially relevant for this exercise, since there are many ways to express the same SQL query (in particular, with different variable names or with a different order of the conditions in the *WHERE* clause).

Exercise 2 (3 points)

Assume we want to find all female German computer scientists listed in Wikidata and where they studied. Formulate the appropriate SPARQL query on the Wikidata dataset linked on the Wiki (use *grep* to figure out the names of the entities and relations needed for the query). Run that query (without *LIMIT*) using your program from Exercise 1, and report the number of query results and the query time in the table on the Wiki. Your program should accept the SPARQL query from the command line as input. Along with the code, also commit a file containing your SPARQL query and corresponding SQL query to the SVN.

SELECT *?x* *WHERE* { *?x* "cast member" "Julia Lockwood" }

Exercise 3 (3 points)

Repeat Exercise 2 for a SPARQL query of your choice, using *LIMIT 10*. That query should have at least two triples in its WHERE clause, and it should have more than one hundred matches without the limit.

Add your code and the files with the SPARQL queries to a new sub-directory *sheet-12* to our SVN, and make sure that everything runs through without errors on Jenkins. As usual, please briefly tell us about your *experiences* with this exercise sheet.